Comments on Miguel Sebastian's
"Experiential Awareness: Do you prefer *It* to *Me*?"

Kenneth Williford
UT Arlington

I will argue that Miguel Sebastian's interesting paper has two serious problems—serious but perhaps not fatal.[1]

1. Though his references indicate that he probably knows better, Sebastian treats a classic and central debate about the phenomenology of subjective character in a rather cavalier manner at a crucial dialectical moment in the paper.

Though we must all turn over our phenomenological spades at some point or other, Sebastian seems to me (and I stress *seems to me*) to be willing to dismiss a whole tradition of thinking about the phenomenology of subjectivity with one of those "it doesn't look that way to me" phenomenological appeals. On page 7, after quoting Kriegel's plug for what Sebastian calls MSI self-representationalism, Sebastian writes "If phenomenal consciousness is essentially self-involving then Kriegel's neobrentanian condition (a state representing itself) does not suffice for an experience to have subjective character. …It seems to me that a certain form of self is essential for an account of the phenomenal character of experiences; it is phenomenologically manifest that my experiences are somehow experiences of

[1] One should bear in mind, of course, that these putative flaws might be an illusion arising from the necessary brevity of the paper.

mine and not that they represent themselves." Again, it could be that Sebastian simply does not have adequate space here to develop his view, but any serious defense of an egological (or "self-involving" to use Sebastian's preferred terminology) theory of consciousness would have to begin by at least understanding its opposition properly. Every non-egological theorist of consciousness from the Buddhists (many schools in any case) to Hume (though it's a rather complicated exegetical story) to G.E. Moore, Sartre and Aron Gurwitsch have been well aware of the fact that we all spontaneously experience our conscious mental episodes as being *our own*. Moreover, many philosophers who clung to faith in the existence of a substantial self, including Kant, Brentano, and Vladimir Solovyov (though this is not so widely known), were quite willing to, in effect, admit that the phenomenology of "mineness" is not great evidence for its existence. Now, I do not think that Sebastian harbors a secret hope that he is an immaterial substance. Rather, the take-home message is twofold:

(A) A lot of phenomenologically acute philosophers over a number of centuries have thought that the phenomenology of "mineness" was perfectly compatible with the non-existence of a substantial self (some thought there is one; some thought there isn't). Moreover, many of the philosophers in question were by no means proponents of the view that the phenomenological data not veridical—on the contrary. They accepted the legitimacy of the phenomenology of "mineness" but held that that was compatible with there being no substantial self (of any kind). Upshot: the substantial self is, plausibly, not given phenomenologically. Upshot for

this paper: It is tendentious and superficial, given the classic nature of exactly this debate, to say "Oh yes it is!" and leave it at that.

(B) But my point goes deeper. Some of these philosophers (e.g., Sartre) articulated fairly elaborate theories that aim at showing how our sense of an autobiographical self emerges on the basis of the primordial reflexivity (prereflective self-consciousness, non-positional self-consciousness, the tacit cogito, etc.) and temporality that characterize all consciousness. In other words, "mineness", on a view like Sartre's, is just identified with the self-manifesting or reflexive (not to be confused with 'reflective') structure of consciousness. Every episode or phase of consciousness—whatever weird story about the individuation of episodes turns out to be true—has this feature. In virtue of that, it is *appearance-to*; it is "for-itself", involves a "dative of manifestation". But even if episodes are metaphysically discrete, they do not seem to be. At the phenomenolgical level, they seem to be retending just prior episodes and protending future ones and *enduring* for a certain temporal span—long enough to observe change. Now, that kernel of "for-itself-ness" remains invariant. This structural universal, the instantiation of which just *is* subjectivity, on this view, characterizes all episodes. When you look down the memory tunnel and you think "this happened to me", that is what you are noticing. It is not as if "mineness" is a special flavor that you can detect when you consider a past experience. You don't know it's yours because you see it has that special character of Ken-Willifordacity or Miguel-Sebastianheit, etc. There no little haecceity you see instantiated by all your experiences; and there's no particular physical object that remains invariant either—only certain structures do. Mineness

is *formal* not *material*—if I may use those terms in their classical senses. On the basis of this formal feature, which characterizes every episode or phase of consciousness, every experience you remember will seem to be yours; and every experience you successfully project will seem to be yours. Over time, after you've abstracted enough of the dispositional profile that characterizes the system in which your stream of consciousness is instantiated, you'll develop a thumbnail sketch of yourself (one that is likely to be quite inaccurate) and start to advertise it. When it is said that Sartre had a "non-egological" theory of consciousness, what is meant is essentially this: the self-manifesting (self-representing, if you like) character of consciousness is prior to any more articulated conceptions of the self. All of those conceptions arise, grow, develop, change (sometimes radically), decay and can dissipate entirely. But one can be conscious (and thus be a "for-itself") throughout all of this. Now, it is not clear, in the final analysis, that Sebastian was wanting to argue that the sort of self involved in his account was what Sartre would call an ego, just as it was not clear that he had some sort of thinking substance in mind. To the extent that he is clear, it seems that he has the self *qua* organism in mind, but we get ahead of ourselves. The main point here is that Sebastian simply ignores the fact that MSI theorists (and here I'd count both Sartre and Kriegel) think that you can build up richer notions of the self (notions that can then be spontaneously and habitually deployed in experience) on the basis of the fundamental reflexivity of consciousness. Sebastian points out that one seems to be aware of oneself in a more robust sense and not merely to be aware of one's mental state as such; and then he moves right on. But this is to just ignore the fact that some "MSI" theorists explicitly

developed accounts of the sort of the very self-awareness he is appealing to. They just do not take that sort of self-awareness to be fundamental. And there is good reason *not* to take it to be fundamental: it changes radically over time, it arguably absent in very young children, in a lot of animals, in late stage Alzheimer's patients, in people having really intense Salvia Divinorum or Datura trips, etc. Again, this is not to say that Sebastian would not offer a way to accommodate these variations in self-representation. But the MSI theorist (like Sartre) has a ready way to do it. Finally, the force of Sebastian's phenomenological appeal can be lessened if we remember that all of these sort of atomistic notions like "state" and "act" (Brentano) tend to make one forget about the temporality of consciousness. They are simplifying but possibly vicious abstractions. If I think of consciousness as an intrinsically dynamic process wherein one episode or phase has arisen from (and perhaps partly overlaps) the previous episode and is in the process of developing into the next phase (which it will partly overlap), then I will understand the phenomenological claim that every episode of consciousness is an awareness of itself much differently from how I would understand the claim if I take it be about some sequence of discrete states that are, in themselves, essentially frozen.

2. I've argued that Sebastian does not take his opposition seriously enough. On the opposition's view, the most basic sense of self (or Core Self, to use Damasio's terms) is provided by the reflexive or self-manifesting structure of consciousness. Once we add more content and history to a conscious system, the sense of an autobiographical self can emerge (and evolve, expand, contract, and ultimately

vanish—get ready for it); but the reflexive structure of consciousness remains invariant. Among many other things, this view is nice because it purports to enable us to capture all of the phenomenologically salient features of subjectivity without our having to postulate some sort of homunculus or substantial self. Now, it becomes clear that Sebastian is not ultimately aiming at reinstating an immaterial, substantial self. While his quasi-formal semantics of *de se* content would certainly be compatible with a Cartesian reading of the individuals that go into his operative ordered pairs, his Damasio-inspired account indicates that it is the material organism that is to play that role.

Now, it's not surprising to me that he picked Damasio's account as a basis here. One common reading of Damasio's account is quite in line with Sebastian's account (and also has some affinities with Thomas Metzinger's account in *Being No One*). I used to call this the "ego-object" version of self-representationalism. On this view, the conscious episodes are just those that represent some privileged ego-object (in this case, the organism) as relating to some other object (of perception, avoidance, attraction, and so on). And yes, fine, sure, couple this with a teleosemantic theory of representation and declare it a Glorious Victory for Naturalism! The intractable problems of teleosemantics (at least vis-à-vis consciousness) aside for the moment, such theories have a deeper problem to surmount.

This is all rather ironic because the problem I am about to articulate is a problem that Dan Zahavi used to accuse neo-Brentanians like Kriegel and me of having. And it is all the more ironic because the problem is forcefully laid out in

Shoemaker's 1968 paper, "Self-Reference and Self-Awareness" to which Sebastian refers. To set up the problem for this context, let's assume we are serious about developing models that really capture the salient phenomenological features of subjectivity. On this assumption, some of the value of a model will be a function of how many features it captures in a "natural" and "coherent" or "unified" way. That's a bit vague, but what I mean can perhaps be said this way: If your model has a primitive where, ideally, it should provide a definition, then it's deficient. If it postulates where it should derive, then likewise. If it generates a conceptual problem where there isn't one in the data it is supposed to represent, then it's really deficient indeed. It can be hard to answer these questions in many cases, but I think that ego-object versions of self-representationalism (or SI FO theories) are deficient in the last way; and one could learn this from Shoemaker and Zahavi (and, perhaps, from Perry, Lewis, and Castañeda as well).

So here is the problem: According to ego-object (SI FO) views, the representational content of conscious mental states is (ultimately) that something is interacting with the privileged object (which is, in effect, the subject-object). In this case the privileged object is the organism. So, we are to imagine that our nervous system has a representation of a good chunk of our whole organized bodily system (something like Damasio's Proto-Self) and that by representing that as being affected by some other object (or whatever variations on that theme you want), *somehow* we get consciousness; and *in particular* we get a sense of self and ownership that colors all our conscious experiences. Now, I see no reason to doubt the existence of the Proto-Self and organism representation, etc. But here is the real

issue: It does not seem like *anything* could make *me* be aware of *this very experience of mine unfolding for me right now* by representing some system (that just happens to be mine) interacting with some other object. The way this relates to Shoemaker's argument presented in the 1968 paper is this: somehow I'd need to know, wouldn't I, that that "organism" represented *is me* or *is mine*! But how, barring a regress? If we are serious about the phenomenology here, that's what the model needs to capture, mineness . But it doesn't! It just says there is this special object (the organism=subject) that is represented along with some interaction with other objects. But that's patently *not* a model of self-awareness! That's a model of some third thing representing two other things interacting. How do you get from that to self-awareness—in *any sense*? Do we just add this: "Oh and, by the way, this organism represented in your brain, *that's you!*" ? Given that there are many ways that I can, like Ernst Mach on the Omnibus, represent myself *without knowing it*, one would think that a model of representing oneself *and knowing it* would not to solve that little problem by fiat or implicit stipulation.

When I teach this stuff I often mention those first-person video games like (sorry for you up-to-date gamers) Quake in which there is a little face down at the bottom of the screen (which otherwise represents one's own visual field). That little face gets bloodier as you get shot or bludgeoned by demons or whatever. But otherwise it is an invariant little object that follows you around throughout all the changes. Now, I submit that basic self-consciousness is *not* a matter of representing some invariant object like that little face. You have to know that that thing is supposed to be you. And to know that, you'd better know who you are already. If I

always saw the floaters in my right eye I would not thereby come to think that they might just be me. And even if I was odd enough to wonder that seriously, I'd still need a more fundamental notion of my own selfhood that was not derived from any such identity judgment—on pain of regress (as Shoemaker and Zahavi rightly point out). And if you can't get basic, primitive selfhood from a *conscious* system making such judgments, then you can't get it out of an unconscious one doing something analogous. It's a complete non-starter.

To sum up: Sebastian too easily dismisses MSI (non-egological) views on the basis of phenomenological considerations that have not at all escaped non-egological theorists of consciousness; on the contrary, some of them have taken those considerations into detailed account. And Sebastian's own account of the representational content of consciousness seems, by some extrapolation, to have been refuted by Sydney Shoemaker in 1968. Again, it is hard to see how one system representing two other systems or objects interacting could give rise to (or constitute) a sense of self. But it's not that hard to see how a self-manifesting stream of consciousness could, over time, ground more elaborate conceptions of the self—though we still do not know what self-manifestation is or how it arises in the brain. Let's not fool ourselves about that either.